加权最小二乘法与卡尔曼滤波实时稳像技术

谷 乐，陈志云†

(华东师范大学 计算中心，上海 200062)

摘 要：无人机和车辆行驶等情况下拍摄的视频，受外界影响会造成视频抖动。通过对比现有的电子稳像技术，提出了利用 FAST 获取特征点的位置信息，再通过光流法结合 NCC 匹配得到参考帧特征点在当前帧的位置信息，在此基础上，结合 RANSAC 算法剔除错误匹配的特征点对的改进算法。为了提高运动矢量估计的精度，应用加权最小二乘法得到相邻帧间的刚性变换矩阵，并经过卡尔曼滤波进行运动平滑得到扫描运动矢量并补偿，最终得到实时的稳定视频。实验表明，视频序列稳像后的帧间变换保真度有所提高，并且能够达到实时处理速度。

关键词：电子稳像技术；特征点匹配；最小二乘法；卡尔曼滤波；运动补偿

中图分类号：TP391.41 doi: 10.3969/j.issn.1001-3695.2018.05.0355

Improved least-squares and Kalman filtering real-time image stabilization

Gu Le, Chen Zhiyun†

(*East China Normal University*, *Shanghai* 200062, *China*)

Abstract: Video taken in the case of drones and vehicles traveling would be affected by the outside world causing video jitter. This paper proposes to use FAST to obtain the position information of feature points by comparing with the existing electronic image stabilization technology. Then the position information of the reference frame feature point in the current frame is obtained by the optical flow method combined with NCC matching, based on this, combined with RANSAC algorithm to eliminate the wrong matching feature points pairs. In order to improve the accuracy of motion vector estimation, this paper applies weighted least squares method to obtain the rigid transformation matrix between adjacent frames. The video is smoothed by Kalman filter to get the motion vector and compensated, and finally get a stable video in real time. The experimental table shows that that the fidelity of the inter-frame transform after video sequence stabilization is improved, and the real-time processing speed can be achieved.

Key words: Electronic image stabilization; feature points matching; least squares method; Kalman filtering; Motion compensation

0 引言

在车载摄像机或者小型无人机拍摄出来的视频中，由于平台不能一直保持很稳定的状态，就会出现视频抖动的情况。消除视频中的抖动，提高视频的质量和观看感受是视频稳定技术研究的重点。目前国内外防抖技术的研究方法包括了机械、光学、机电以及电子防抖等[1]。对比于其他方法，电子稳像技术因具有成本低廉，对于硬件设备要求不高，而且稳像效果很好的优势，成为一种广受关注的研究对象。

电子稳像技术首先利用运动矢量估计得到运动补偿参数，然后平滑补偿参数，最后将平滑后的补偿参数对原始抖动序列进行补偿，从而得到稳定的视频。其中，运动矢量的计算的准确度对整体算法的影响最大，因此如何精确地获取运动矢量变化信息是电子稳像方法中亟待解决的核心问题[2]。运动矢量估计算法包括了灰度投影法[3]、位平面匹配法[4]、代表点匹配法[5]、特征点匹配法[6]等。

本文比较了不同算法的优劣性，SIFT（scale invariant feature transform）方法被广泛认为是当前最优秀的角点探测器之一，其提取到的特征量极具辨识力，之后出现的一些算法都是保留 SIFT 方法有效性的基础上优化运算时间，其中，SURF(Speeded up robust features)被认为是最成功的改进方案[2]。但是 SURF 计算复杂度很高，很难做到实时应用。FAST（features from accelerated segment test）是由 Rosten 和 Drummond 于 2006 年提出一种简单、快速的角点检测算法，比 SIFT 和 SURF 方法都要快很多倍[7]。本文采用 FAST 特征点检测进行帧间配准，转换得到帧间局部运动矢量，利用光流法结合 NCC(normalized

收稿日期：2018-05-11；**修回日期：**2018-07-04

作者简介：谷乐（1991-），女，黑龙江泰来人，硕士研究生，主要研究方向为教育技术、图像处理；陈志云（1967-），女（通信作者），上海人，副教授，主要研究方向为多媒体技术（chenzhy@cc.ecnu.edu.cn）.

cross correlation)匹配获得参考帧特征点在当前帧的位置信息。通过利用 RANSAC(random sample consensus)算法剔除掉“坏”的特征点对，然后应用改进的最小二乘法得到相邻帧间的刚性变换矩阵，并经过卡尔曼滤波和双线性插值法进行双线性补偿最终得到实时稳定的视频。

1　特征点检测和表示算法

图像匹配在基于特征匹配的电子稳像技术中是非常关键的，本文给出了实时数字视频稳像方案，如图 1 所示。

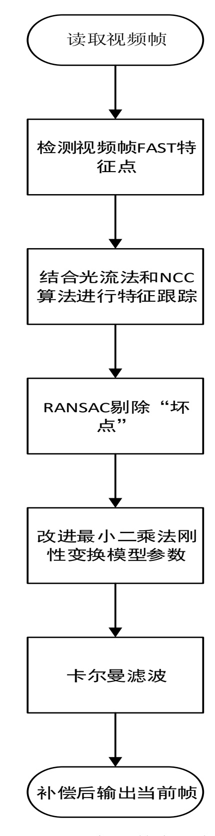


图 1　基于特征点匹配的实时数字视频稳像方法流程图

1.1　FAST 特征点检测

FAST 全称是 features from accelerated segment test，该算法的原理是从图像中取一个检测点，以这个点为圆心利用周围 16 个像素点来判断该点是否为角点。如图 2 所示，先判断 1、5、9 和 13 这 4 个像元的灰度值能否满足条件，满足则继续判断，不能满足就剔除该点。利用该算法平均检测角点周围 3.8 个点就可以判断出检测点是否为角点。表 1 列举了 FAST、SIFT、SURF 这三种检测算法的比较。

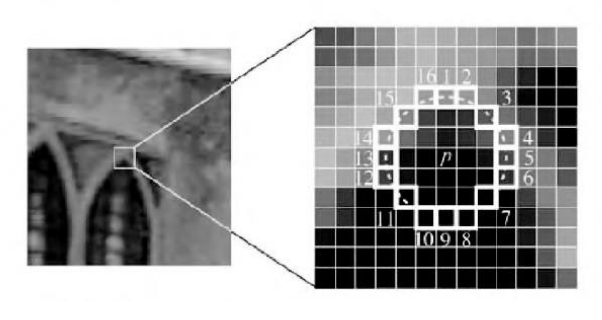


图 2　FAST 特征点检测示意图

表 1　常用特征点检测算法的比较

检测方法	特征点数/帧	平均检测时间/ms/帧	平均每个特征点检测时间/ms
FAST	300.501	0.583	0.0019
SIFT	300.320	385.357	1.2832
SURF	299.899	96.291	0.3208

从表 1 可以看出， SIFT 和 SURF 特征提取算法比 FAST 计算量大很多，耗时也很长。从算法效率角度考虑，FAST 具有比较明显的优势。图 3 给出了 FAST 算子提取到的特征点示意图。



图 3　FAST 提取特征点示意图

1.2　基于金子塔 Lucas-Kanade 与 NCC 结合的特征点匹配

利用特征点检测得出特征点位置信息后，还需要算出当前帧的特征点的位置信息。本文利用金子塔 Lucas-Kanade 光流算法计算。

Lucas-Kanade(L-K)算法[8]有 3 个假设前提：a)亮度保持恒定;b)相邻帧间的运动随时间变化缓慢或者相邻帧在时间上连续;c)保持空间一致性。

Lucas-Kanade(L-K)算法效率高，实时性较好，对于较为突出的特征点能够进行有效地跟踪，但是不能够很好地适应光照的变化。而且，该算法在跟踪不是很突出的特征点时容易发生错误。归一化交叉相关算法（normalized cross correlation）[9]可以弥补 L-K 的不足。首先 NCC 可以有效降低光照对图像比较结果的影响，此外，该算法比较容易量化比较结果，而且该算法通过调整窗口的大小和每次检测的步长矩形，在一定程度上提高了跟踪效率和准确率。在工业生产环节检测、监控领域对像检测与识别等场景下比较两幅图像的相似程度通常会用到 NCC 算法。

本文在进行特征跟踪时利用 L-K 和 NCC 算法结合。对选取的每个特征点按照响应值大小进行划分，比阈值大的点采用 NCC 算法进行跟踪，比阈值小的点使用 L-K 算法进行跟踪。NCC 的计算代价可以使用积分图像方法优化归一化计算过程和使用频域乘积优化相关（卷积）计算过程[10]。为了减小计算代价，同时采用金字塔策略。图 4 展示了第 306 帧视频特征点追踪的效果。

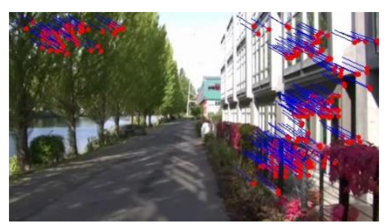


图 4　第 306 帧视频特征点追踪

1.3　RANSAC 算法

为了得到更为精准的全局运动矢量，防止随机运动的物体上的特征点局部运动影响全局相对运动矢量估计，本文采用 RANSAC 算法[11]剔除运动目标的干扰。

经典的 RANSAC 算法由 Fischer 和 Bolles[12]在 1981 年首先提出。经典 RANSAC 是不断地尝试不同的目标空间参数，最大化目标函数的过程。这个过程中需要随机地选择数据集的子空间，然后能够产生一个模型估计，使用数据集剩余的点进行测试估计出来的模型，获得一个得分，最终得分最高的模型估计就是整个数据集的模型。而判断当前某个点是否为类内则需要设定一个阈值。

每一个子集应是在迭代的过程中一个大小为 m 的最小采样。最小采样就是 m 的大小满足计算 θ 的个数。至少有一次在循环过程中的采样，在置信度为 η_0 的条件下能够使采样出的 m 个点均为类内点，只有这样才能够保证至少有一次采样能在循环的过程中，取得目标函数的最大值。因此采样次数 k 应该满足的条件为

$$k \geq \frac{\log(1-\eta_0)}{\log(1-\varepsilon^m)} \tag{1}$$

其中：m 代表子集大小，ε 代表类内点的比例，置信度一般设置在成为[0.95,0.99]。循环终止的条件是将选取的子集是“全部是类内点”或“不全都是类内点”这两种结果的二项分布，“全部都是类内点”的概率为 $p=1/\varepsilon^m$。如果 p 足够小，就将其作为一种泊松分布。因此，有 n 个“子集全部都是类内点”的概率表达为

$$p(n,\lambda)=\frac{\lambda^n e^{-\lambda}}{n!} \tag{2}$$

λ 表示“子集全部都是类内点”的选取次数的期望。

2　基于改进最小二乘法刚性变换模型参数估计

2.1　传统的运动矢量估计

经过 RANSAC 筛选后，采用 4 参数的仿射变换模型[12]计算运动矢量：

$$\begin{pmatrix} X_n \\ Y_n \\ 1 \end{pmatrix} = \begin{pmatrix} a & -b & t_x \\ b & a & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_{n-1} \\ Y_{n-1} \\ 1 \end{pmatrix} \tag{3}$$

其中：a 代表图像的缩放分量，b 代表旋转分量，t_x 和 t_y 代表图像间平移分量，(X_{n-1},Y_{n-1}) 代表参考帧的特征点坐标，(X_n,Y_n) 代表当前帧的特征点坐标。相邻帧间的刚性变换矩阵可以通过最小二乘法得到。

刚性变换可以表示为

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_k \\ y_k \end{pmatrix} \tag{4}$$

2.2　改进的运动矢量估计

如果采用传统的最小二乘法就只能大致计算出运动矢量，因为不同的特征点景深会严重影响运动矢量的准确性[13]。所以本文对传统的最小二乘法进行改进，使用粒子滤波的方法得到特征点权重，再通过使用加权最小二乘法得到更加准确的运动矢量。粒子滤波的基本思想[14]是利用带权重的粒子 $\{X_t^i,\omega_{t-1}^i\}_{i=1}^N$ 近似获得系统状态的后验概率密度 $p(X_{t-1}|Z_{1:t-1})$，并估计系统状态。其中 X_t 和 Z_t 分别为目标状和 t 时刻的观测值。

首先随机选用一组特征点对组成粒子，然后利用式（3）得出运动矢量，需要注意的是，特征点对要多于 3 对。一个特征点可能属于多个粒子，获得 N 个粒子后，它的权重由逐点均方准则确定，如式（5）所示。

$$\omega_t^i = \frac{\exp\left[-\frac{(I_{i-1}-I_t^i)^2}{2\sigma_t}\right]}{\sum_{i=1}^{N}\exp\left[-\frac{(I_{t-1}-I_t^i)^2}{2\sigma_t}\right]} \tag{5}$$

其中：ω_t^i 表示在 t 时刻第 i 个粒子的权重；I_{t-1} 为 $t-1$ 时刻的图像；σ_t 代表方差。I_t^i 是 I_t 利用式（5）基于第 i 个粒子计算的第 i 帧补偿图像。因为补偿图像会存在无定义区，所以在计算的时候取参考帧的中心区域 $\omega/2*h/2$，ω 代表参考帧图像的宽度，h 代表参考帧的高度。通过粒子的权重就能获得特征点的权重，即

$$W_t^j = \frac{1}{N_t^j}\sum_{k=1}^{N_t^j}\omega_t^k \tag{6}$$

其中：W_t^j 是在 t 时刻第 j 个特征点的权重；N_t^j 是第 j 个特征点所对应的粒子数；ω_t^k 是相应粒子的权重。然后就可以运用加权最小二乘法准确地估计当前帧的运动矢量。

$$M = \left[X^T \Lambda X\right]^{-1} X^T \Lambda Y \tag{7}$$

其中：$\Lambda = diag(W_t^1, W_t^2, \ldots W_t^N)$，$N$ 是在一帧图像当中的所有特征点的个数。在选择特征点数目时同时要考虑到，每个粒子的特征点数目不相同，就会影响到全局运动矢量的准确。

3　卡尔曼运动补偿与视频输出

卡尔曼滤波用来从运动矢量中提取出有意运动矢量，整个过程是根据前一时刻的估计值和协方差估计值预测出当前时刻的运动矢量。卡尔曼滤波的预测方程为

$$\begin{cases} s(k|k-1) = Fs(k-1|k-1) \\ p(k|k-1) = Fp(k-1|k-1)F^T + Q \end{cases} \tag{8}$$

卡尔曼滤波器的更新方程为

$$\begin{aligned} &S(k|k) = s(k|k-1) + k_g(k)[z(k) - Hs(k|k-1)] \\ &P(k|k) = (I - k_g(k)H)P(k|k-1) \\ &k_g(k) = HP(k-1|k-1)H^T(HP(k|k-1)H^T + R)^{-1} \end{aligned} \tag{9}$$

式(8)中：$s(k)$ 为系统的状态矢量，F 代表状态矩，$p(k)$ 为相对应的协方差矩阵，F^T 为 F 的转置矩阵，Q 为测量噪声的方差；式(9)中，$S(k)$ 就是卡尔曼滤波的结果，k_g 为卡尔曼增益，H 代表观测矩阵，$P(k)$ 为滤波后的协方差。Q、R 通常取固定值。

4　稳像实验验证

本文选取了几组室外拍摄的抖动视频，先通过部分视频的稳像前后对比图来主观的评价稳像效果，实验表明，在进行帧图片对比时，不同的稳像算法差别不明显，但处理连续多帧视频时，时间差异和视频质量上会产生差异，就需要通过峰值信噪比来客观的比较稳像效果。本文实验所用的视频分辨率480*270 的视频。

4.1　主观验证

本文截取了连续的 3 帧视频显示稳像前后的效果，如图 5 所示。



(a)稳像前



(b)稳像后

图 5　稳像前、后效果图对比

图 5 中，（a）为稳像前的图片，（b）为稳像后的图片，以窗户标记处为基准画两条白色平行线观察。可以发现，稳像后图片中的景物基本保持在同一水平线上。整个视频处理过后，明显比处理前平稳，减轻视觉疲劳，提高了视频的观赏性。再通过用不同方法处理同一视频，从视觉效果上，大部分方法都有一定的稳定的效果，帧图片上的效果差别不明显。

4.2　PSNR 值验证

PSNR（peak signal to noise ratio）作为评价稳定算法标准度的指标，用来衡量两幅图像重合的情况，PSNR 值越大，图像稳定效果越好[14]。其计算公式如下：

$$PSNR = 10\log_{10}\left(\frac{(2^n-1)^2}{MSE}\right) \tag{10}$$

$$MSE = \sqrt{\frac{\sum_{x=1,y=1}^{M,N}\left(I_{cur}(x,y) - I_{prep}(x,y)\right)^2}{M\times N}} \tag{11}$$

式(11)用来计算相邻帧之间像素值的均方误差 MSE，M 和 N 分别为图像的长和宽，由于图像灰度范围为 0~255，所以，本文中 n 为 8。$I_{cur(x,y)}$ 表示当前帧图像灰度值，$I_{perp(x,y)}$ 表示前一帧图像灰度值。Song’s[7]算法采用了 SURF 算子和加权滤波的方法，也是目前电子稳像算法中比较有代表性的算法。图 6 表示出了稳像前、Song’s 算法、本文算法对于从第 10 帧到第 30 帧后的 PSNR 值。图 7 表示出了对比方法和本文方法从第 10 帧到第 30 帧的稳像时间的对比。对比方法中采用了 SURF 算子，由图 6 可以看出，本文方法的 PSNR 值比稳像前和对比方法在大部分视频帧画面中都高，而且处理速度比比较方法更快。经实验证明，本文的方法更适用于于实时稳像。

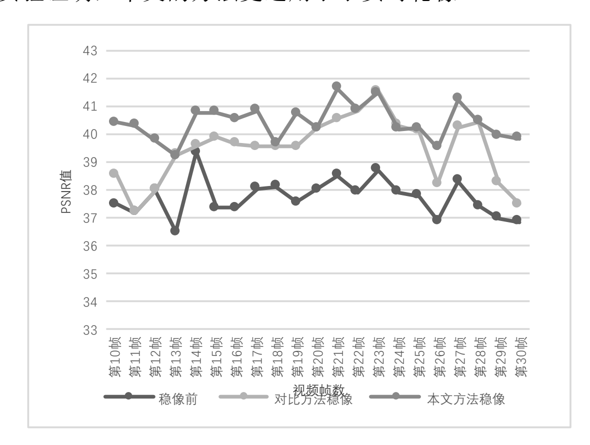


图 6 稳像前、对比方法稳像、本文方法稳像后 PSNR 对比图

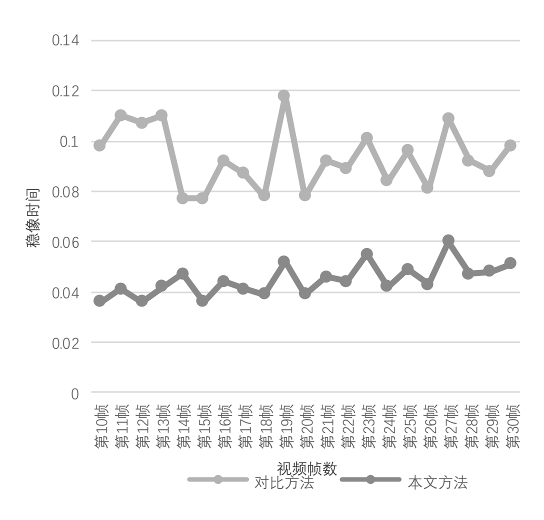


图 7　对比方法稳像、本文方法稳像时间对比图

5　结束语

本文采用 FAST 特征匹配的方法，利用金字塔 L-K 与 NCC 结合的方法对特征点进行追踪，通过 RANSAC 算法剔除错误的匹配点，提高运动估计矢量精度，利用加权最小二乘的粒子滤波方法对运动矢量进行增强，改进了运动矢量估计过程。接着采用卡尔曼滤波得到抖动分量进行补偿，得到了效果较好的视频效果。实验结果表明本文设计的稳像方法可以明显地去除抖动，稳定图像，并且在处理时间上具有优势，峰值信噪比也有所提高。下一步工作是进一步提高稳像后视频清晰度和减小黑边。

参考文献：

[1] 唐佳林，郑杰锋，李熙莹．基于特征匹配与运动补偿的视频稳像算法[J]．计算机应用研究, 2018, 35 (2): 608-614. (Tang Jialin, Zheng Jiefeng,

Li Xiying. Video stabilization algorithm based on feature matiching and motion compensation [J]. Application Research of Computers, 2018, 35 (2): 608-610.)

[2] 熊晶莹. 基于特征提取与匹配的车载电子稳像方法研究 [D]. 北京：中国科学院大学，2017. (Xiong Jingying. Study of electronic image stabilization based on feature detection and description for vehicles [D]. Beijing: University of Chinese Academy of Sciences, 2017.)

[3] Puglisi G, Battisto S. Fast block based local motion estimation for video stabilization [C]// Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2011: 50-57.

[4] Tian Shaoqing, Zhao Peng, Wang Nihong, *et al.* Aims at moving objects'improvement based on gray projection of algorithm of the electronic image stabilization [C]// Proc of the 3rd International Congress on Image and Signal Processing. 2010: 2483-2487.

[5] Natarajan B, Bhaskaran V, Konstantinides K. Low-complexity block-based motion estimation via one-bit transforms [J]. IEEE Trans on Circuits and Systems for Video Technology, 1997, 7 (4): 702-706.

[6] Tang Jin, Han Xiaowei, Yuan Zhonghu. An approach of electronic image stabilization based on the representative point matching [C]// Proc of the 3rd International Conference on Genetic and Evolutionary Computing. 2009: 345-350.

[7] Song Chunhe, Zhao Hai, Wei Jing, *et al.* Robust video stabilization based on particle filtering with weighted feature points [J]. IEEE Trans on Consumer Electronis, 2012, (58): 570-577.

[8] 聂婷. 基于改进 FAST 特征匹配的电子稳像算法 [J]. 长春：中国科学院长春光学精密机械物理与物理研究所，2015: 38-11. (Nie Ting. Electronic image stabilization algorithm based on improved FAST feature matching [J]. Changchun: Changchun institute of optics Fine Mechanics and Physics, Chinese Academy of Sciences, 2015: 38-11)

[9] Bruce D L, Takeo K. An interactive image registration technique with an application to stereo vision [C]// Proc of the 7th International Joint Conference on Artificial Intelligence. 1981: 2: 674-679.

[10] 董晶，杨夏. 基于单应轨迹的视频实时稳像算法 [J]. 国防科技大学学报，2014 (36): 99-104. (Dong Jing, Yang Xia, Yu Qifeng, A real-time video stabilization algorithm based on homography trajectory [J] , Journal of National University of Defense Technology, 2014 (36): 99-104.

[11] Hii A J H, Hann C E, Chase J G, *et al.* Fast normalized cross correlation for motion tracking using basis functions [J]. Computer Methods and Programs in Biomedicine, 2006, 82 (2): 144-156.

[12] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24: 381-395.

[13] Walha A, Wali A, Alimi A M. Video stabilization for aerial video surveillance [C]// Proc of AASRI Conference on Intelligent System and Control, 2013, 4: 72-77.

[14] 逯晓光，靳万鑫，余颖. 大运动前景和旋转抖动视频的快速数字稳定 [J]. 光学精密工程，2014, 22 (1): 178-185. (Di Xiaoguang, Jin Wanxin, Yu Ying. Digital video stabilization for large-scale moving foreground object and rotation jitter [J]. Optics and Precision Engineering, 2014, 22 (1): 178-185.)